Database Technology Professional (DBTech Pro) Project

WP 2 – Specification of Knowledge Areas

Final report:

Thomas M Connolly, Work Package Leader
University of Paisley
School of Computing
High St, Paisley, PA1 2BE
Scotland
email: thomas.connolly@paisley.ac.uk

Table of Contents

1 Introduction

The database is now the underlying framework of the information system and has fundamentally changed the way many organisations and individuals work. This is reflected within tertiary education where databases form a core area of study in undergraduate and postgraduate programmes related to computer science and information systems, and typically at least an elective on other data-intensive programmes (ACM/IEEE 2001; EUCIP 2003). The core studies are commonly based on the relational data model, SQL (the *de facto* language for relational DBMSs), data modelling and relational database design. This curriculum initially supported industry needs where the relational DBMS was the dominant data-processing software. At the same time, the field of computing has seen significant developments over the last decade with object-orientation becoming much more predominant and the Web emerging as a global communication and business medium. Database technology has encompassed these new developments with the emergence of object-relational DBMSs, object-oriented DBMSs, support for Web-database integration, HTML, XML, XQuery, Java, J2EE, data warehousing, business intelligence, and so on. These developments have significantly increased the breadth and depth of knowledge that a modern database professional requires to work effectively in industry.

The Database Technology Professional (DBTech Pro) project is an EU, Leonardo da Vinci Program Pilot Project involving the participation of academic and industrial partners from the Database Technology Network (DBTechNet) initiative (http://www.dbtechnet.org) that is investigating how academia can better support the needs of the database community. The partnership extends across five EU member states (Finland, Germany, Greece, Spain, UK) and involves seven academic institutions: Reutlingen University, Germany; Häme Polytechnic, Hämeenlinna, Finland; Helia (Helsinki Business Polytechnic), Helsinki, Finland; TEI of Thessaloniki, Greece; University of Macedonia, Thessaloniki, Greece; University of Malaga, Spain; University of Paisley, UK; and three companies: ALTEC S.A., Thessaloniki, Greece; Solid EMEA North Finland; TietoEnator Public Sector Finland. The project aims include:

(a)   the identification of professional roles and skills in relation to contemporary database technology practice in Europe;

(b)   for each one professional role identified, the specification of the knowledge and skills that are needed by the individual who wishes to achieve competency in the relevant job market;

(c)   the development of pilot course and laboratory workshop material for training and, subsequently, testing the knowledge and skills in question;

(d)   an investigation into the possibility of developing a vendor-independent European certificate of Database Technology Professional.

In this respect, the goals of the DBTech Pro project are along the lines of, and complementary in nature to, those of the broader industry-driven, vendor-independent professional certification and competency development initiatives (e.g. EUCIP) that target the database professional

and practitioner. Direct beneficiaries of the project include:

- Students studying their degree in the partner institutions will obtain basic knowledge and skills as a database professional for employment opportunities both with national and international enterprises.
- Enterprises will be assisted to catch-up with the rapid pace of developments in database technology, assess and adapt to the existing databases. In this respect, they will avoid pitfalls and they can employ highly trained and competitive staff. The proposed European DBTech Pro Certificate for a database expert will guarantee the quality of knowledge and the competence in modern database technology of the employees. In a larger scale this will benefit the competitiveness of the whole European industry on worldwide markets, since professional use of databases is the foundation of all information systems.
- Database professionals in SMEs having the proposed European DBTech Pro Certificate can prove their knowledge and skills – and their value in the labour market.
- Educational institutions will combine their human and technical resources and thus get better results for all their staff members and students.

To enable the efficient and effective organisation of the collaboration across the partner institutions, the project has been decomposed into a number of work packages:

WP1: Project Management
– to oversee the effective and efficient organisation of the project.

WP2: Specification of Knowledge Areas 5
– to determine the knowledge areas required by a European Database Technology professional.

WP3: Specification of Teaching Practices and Laboratories, Pilot Tests
– to determine what teaching modules and laboratories exist across the partners, and to identify and develop four pilot workshops during the project.

WP4: Content Planning
– for each the knowledge areas identified in WP2, compile an initial version of educational/course content outline, indicative of the database technology knowledge and skills to be assessed for professional competence in the job market.

WP5: Pilot Run of the DBTech Pro
– to establish pilot runs of the four workshops that will be created during this project among the partner organisations.

WP6: Assembly of the DBTech Pro Certificate
- – to make recommendations on the industry support for a vendor-independent certification scheme for a European Database Technology professional.

WP7: Evaluation and Assessment
- – to evaluate the project overall and the four pilot workshops to be created during the project.

WP8: Dissemination
- – on an on-going basis, to disseminate the results of the project.

WP9: Workshops Planning and Organisation
- – to plan and organise the four pilot workshops to be developed during this project.

2 Work Package 2 – Specification of Knowledge Areas

The project required a robust solution that would identify both contemporary and future knowledge areas. It was therefore decided to use a survey-based approach that would seek to identify the current roles that were in use in industry within the partner countries but would also seek to identify the database skills that companies thought were lacking in both recent graduates and employees and their plans for future database development. At the same time, the survey would also seek to determine the interest for a vendor-independent certification scheme for a European Database Technology professional that would provide data for input to WP6.

The aim of this particular report is to discuss the methodology used in the partner countries for this survey and to present the main findings of the survey. The survey is provided an appendix to this document.

3 Methodology

The survey was designed as a collaborative effort by the entire team and an online version of the survey was implemented on the DBTech Pro web site. The survey was carried out in each country using a slightly different approach that took account of country differences. In the **UK**, the West of Scotland was used as a representative region of the UK and this region then used for sampling. During the period from April to August 2003 a total of 212 surveys were sent out to a representative sample to cover all business sectors and 70 returns were received.

To define the relevant target group in **Finland**, the team asked:
- The R&D manager of FIPA (Finnish Information Processing Association) to send a

message to the corporate members of FIPA and ask the companies to nominate the respondents. There are only 263 companies that have greater than 500 employees. FIPA has 700 corporate members. The R&D manager picked up 109 of these to whom the request to response was sent.

- Some selected vendor companies (ORACLE, IBM, MS, Sybase and Solid) that control a major part of the market to send the same request as FIPA to their client companies in Finland. This was not successful because the vendors declined. Instead of the vendors the team contacted the user groups.
- The Rela Club members to respond. Sytyke is a respected Finnish society of software engineering that includes a number of clubs and Rela is one of the clubs. Rela is in practise the only vendor-independent database specialist club in Finland.
- Helia's industrial partners and the so-called 'Helia support group' of circa. 40 companies and other organisations.

For Finland, a total of 52 responses were received.

In **Greece**, the survey was distributed to 150 companies but only 9 responses received.

During the period from July to August 2003 in **Germany** the survey was sent to approximately 300 companies. The addresses were taken from Reutlingen University's industrial contact database, selecting people who are working for IT-companies in any position. Due to the purpose of this database the target people included mainly executive managers, IT-professional, project managers, alumni in IT, and IT-marketing people. 74 addresses out of these were no longer valid (for example, people had moved to a different employer or changed address). As media email was used to ask people to complete the online questionnaire. There were only seven online responses and two people returned the questionnaire via fax. The reason for the low response rate was according to the comments received when telephoning was mostly that they had no time, that they were fed up with surveys, or that they were on holiday.

Finally in **Spain**, since there are no professional communities/groups in Spain, the sampling unit was determined to be the IT enterprises. The sampling frame was a list in an Internet directory of IT enterprises (Spanish enterprises and Spanish subsidiaries). This directory contains up to 15,000 registered enterprises, so only those ones with registered contact email and at least 17 employees were selected. An email was sent to 517 enterprises with a cover letter explaining the objectives of the survey and a link to the online survey on the DBTech Pro web site. Due to server problems, wrong and old emails, etc. it was estimated that 10% of the messages did not reach their final target. Only 3 enterprises completed the questionnaire. It is believed that this was mainly because: (a) DBTech is not yet known in Spain; (b) there was insufficient motivation to complete the questionnaire; (c) the email was viewed as *spam*.

Due to the low response from Greece, Germany and Spain, the analysis was based on the responses from the UK and Finland.

4 Survey Results

4.1 Roles and Numbers

What database roles does your company currently have (please tick as many as apply and indicate the approximate number currently employed in the role together with the current number of vacancies).

Table 1 summarises the returns from this particular question. By a significant margin, the Database Programmer/Developer is the most common role just now, with Data Warehouse Manager and Data Miner both minor roles by comparison.

Q1 – Roles & Numbers		
	Employed	**Vacancies**
DB Administrator	12.1%	7.4%
DB Designer/Analyst	18.0%	13.5%
DB Programmer/Developer	61.7%	74.0%
DW Manager	4.8%	3.0%
Data Miner	3.4%	2.0%

Table 1. Database Roles in Companies

4.2 DBMSs and Platforms

Which DBMSs do you mainly use (please tick those that apply and indicate whether based on NT or Unix/Linux)?

Table 2 summarises the returns from this particular question. Three DBMSs are heavily used by the sample companies (Oracle, Microsoft SQL Server and Microsoft Access) with MySQL fourth with 11.7%.

Q2 - DBMSs and Platforms	
Oracle	25.5%
DB2	8.0%
Ingres	0.0%
Sybase	4.4%
SQLServer	20.4%
Access	21.2%
MySQL	11.7%
Solid	2.2%
Others	6.6%

Table 2. Used DBMSs and Platforms

4.3   Database Application Development

Is you database application development performed in-house, outsourced or neither?

Table 3 summarises the returns from this particular question. Of the companies surveyed, approximately half carry out their database application development in-house, while 15% out-source the development work. Surprisingly, 35.4% neither develop their database applications in-house or out-house, which seems to suggest use of off-the-shelf packages.

Q3 - Database Application Development	
In-House	49.6%
Out-Sourced	15.0%
None	35.4%

Table 3. Database Application Development

4.4   Graduates (Lack of Skills)

Please specify which database skills/knowledge (if any) you feel are lacking in recent graduates you have hired to work on database related jobs.

Table 4 summarises the returns from this particular question. The responses were quite wide-ranging but a number of generic categories were identified. Database Administration, Database Design, and Tuning were common issues.

Q4 - Graduates (Lack of Skills)	
DB Administration	20.2%
DB Design	28.6%
Normalisation	6.5%
Tuning	25.0%
DB Development	12.5%
Others	7.1%

Table 4. Graduates, Lack of Skills

4.5   IT Staff (Lack of Skills)

Please specify which database skills/knowledge (if any) you feel are lacking in your IT professionals working on database related jobs.

Table 5 summarises the returns from this particular question. Once again, the responses were quite wide-ranging but as with the last question a number of generic categories were identified. Database Design and Tuning were common issues with Database Administration the third common lack of skills within existing staff.

Q5 - IT Staff (Lack of Skills)	
DB Administration	16.2%
DB Design	30.3%
Normalisation	10.1%
Tuning	26.3%
DB Development	7.1%
Others	10.1%

Table 5. IT Staff, Lack of Skills

4.6   Languages and Tools Used

Please specify which languages/tools you use for the following.

Backup and Recovery

Table 6(a) summarises the returns from this particular question. Native tools (RMAN, DB2 tools, and other vendor-supplied tools) are the most popular.

Q6 - Languages/Tools		
Backup/Recovery		
RMAN	30	38.5%
NT	13	16.7%
Unix	2	2.6%
DB2 Tools	16	20.5%
BMC Tools	3	3.8%
Native Tools	6	7.7%
Veritas	4	5.1%
SQL	1	1.3%
C	1	1.3%
Shell scripts	1	1.3%
OS related	1	1.3%
Total	78	

Table 6(a). Used Languages / Tools for Backup and Recovery

Database Design

Table 6(b) summarises the returns from this particular question. A number of tools are currently popular for database design, most notably Rational Rose, ERWin, and Oracle Designer.

Q6 - Languages/Tools		
Database Design		
Select SSADM	6	8.2%
Rose	22	30.1%
ERWin	17	23.3%
Oracle Designer	15	20.5%
Sybase PowerDesigner	5	6.8%
Coolgen	1	1.4%
Rochade	1	1.4%
Advantage GEN	1	1.4%
Visual Studio	1	1.4%
SQL	1	1.4%
Visual basic	1	1.4%
UML	2	2.7%
Total	73	

Table 6(b). Used Languages / Tools for Database Design

Database Development

Table 6(c) summarises the returns from this particular question. There are currently three main languages for database development: PL/SQL, Visual Basic, and C++.

Q6 - Languages/Tools		
Database Development		
PL/SQL, SQL	27	34.6%
VB	19	24.4%
C++	15	19.2%
Java	10	12.8%
Oracle Designer	4	5.1%
Toad	3	3.8%
Total	77	

Table 6(c). Used Languages / Tools for Database Development

Application Development

Table 6(d) summarises the returns from this particular question. A number of languages are currently popular for application development, most notably Visual Basic, PL/SQL, Forms/Reports, C++, Java. The last one is particularly interesting, with Java having approximately the same usage now as C++.

Q6 - Languages/Tools		
Application Development		
PL/SQL, SQL	22	12.6%
VB	29	16.6%
Forms/Reports	27	15.4%
C++	32	18.3%
Java	31	17.7%
JavaScript	5	2.9%
PERL	4	2.3%
JDeveloper	2	1.1%
JBuilder	2	1.1%
Cobol	8	4.6%
PL/1	4	2.3%
Websphere	4	2.3%
Advantage Gen	3	1.7%
PowerBuilder	2	1.1%
Total	175	

Table 6(d). Used Languages / Tools for Application Development

4.7   Future Development

Do you have in production or planning any of the following developments (please tick as many as apply).

Table 7 summarises the returns from this particular question. The two top future developments were eCommerce and XML. Next were Data Warehousing and J2EE/ .NET. With hindsight, it may have been useful to have split J2EE and .NET, although we might be able to make some inferences based on platforms and languages used for database development.

Q7 - Future Developments	
Data Warehousing	17.6%
ECommerce	29.7%
ERP/CRM	9.6%
J2EE/.NET	16.7%
XML	25.9%
Others	0.4%

Table 7. Future Developments

4.8 Top 3 Developments

Identify the top three (ranked from 1 to 3) skills that in your opinion are most likely to influence Database Technology (and the corresponding job market) the most in the near future.

Table 8 summarises the returns from this particular question. Each of the responses has initially been divided based on their ranking, with the final column showing an overall weighted ranking. The top future development is the Web by quite some margin, followed by Decision Support (that is, data warehousing), followed in joint third by Object-Orientation (object-relational is included within this category) and Data Distribution. This would be in line with current thinking.

Q8 - Top 3 Developments							Overall
	1		**2**		**3**		**Overall**
Object-Orientation	9	8.3%	16	14.8%	9	8.4%	10.5%
Decision Support	11	10.2%	25	23.1%	16	15.0%	15.3%
Business Intelligence	8	7.4%	6	5.6%	13	12.1%	7.6%
Web	67	62.0%	16	14.8%	11	10.3%	37.7%
Parallel Processing	1	0.9%	3	2.8%	3	2.8%	1.9%
MD/Temporal	1	0.9%	5	4.6%	3	2.8%	2.5%
Deductive	0	0.0%	1	0.9%	2	1.9%	0.6%
Distribution	6	5.6%	14	13.0%	21	19.6%	10.4%
Real-Time	2	1.9%	9	8.3%	11	10.3%	5.4%
Workflow	2	1.9%	1	0.9%	10	9.3%	2.8%
Agents	1	0.9%	12	11.1%	8	7.5%	5.4%
	108		108		107		

Table 8. Top 3 Developments

4.9 Database Certification

Please specify the approximate number of staff who have the following database vendor certification.

Table 9 summarises the returns from this particular question. Five particular certifications were identified in the returns: Oracle, Microsoft, IBM, Sybase, and, surprisingly, MySQL. The first two were the main certification schemes for staff. This represents about 8.7% of all database staff in the companies (which seems quite high).

Q9 - Database Certification	
Oracle	41.6%
MS	33.5%
IBM	19.0%
Sybase	4.5%
MySQL	1.4%

Table 9. Number of Staff with Database Vendor Certification

4.10 Importance of Certification

Do you consider database vendor certification to be important for staff?

Table 10 summarises the returns from this particular question. The data would suggest that industry is not that concerned about the existing certification schemes, with a slight preference for certification.

Q10 - Importance of Certification	
Not Important	27.8%
Neutral	28.7%
Somewhat Important	38.9%
Very Important	4.6%

Table 10. Importance of Certification

4.11   Independent Certification

Would you prefer to see a vendor-independent certification scheme for database professionals?

Table 11 summarises the returns from this particular question. There seems to be clear evidence that there is a desire for an independent certification scheme with more than 80% in agreement and less than 3% in disagreement.

Q11 - Independent Certification	
Strongly Disagree	0.9%
Somewhat Disagree	1.9%
Neutral	15.0%
Somewhat Agree	51.4%
Strongly Agree	30.8%

Table 11. Prefer to See a Vendor - Independent Certification

5 Specification of Knowledge Areas

Amalgamating the results from Questions 1, 4, 5, 7 and 8 together, we believe that the knowledge areas for future development in later work packages (for example, Work Package 4) are: Data Access Technologies (Database Programmer), Database Design, Database Administration, and OLAP & Data Warehousing.

6 Summary

The Database Technology Professional (DBTech Pro) project is an EU, Leonardo da Vinci Program Project involving the participation of academic and industrial partners. The aim of this particular report is to discuss the survey that was performed as part of this project. The survey was designed as a collaborative effort by the entire team and an online version of the survey was implemented on the DBTech Pro web site. The survey was carried out in each country using a slightly different approach that took account of country differences, however, only the surveys in the UK and Finland produced significant responses.

In terms of professional roles currently in use in industry, the survey found that the main roles were Database Programmer/Developer (61.7%), Database Designer/Analyst (18%) and Database Administrator (12.1%). In terms of lack of skills, the three top deficiencies were Database Design (35.1% in graduates and 40.4% in current employees), Tuning (25% in graduates and 26.3% in current employees) and Database Administration (20.2% in graduates

and 16.2% in current employees). The Web and Data Warehousing/Data Mining (Decision Support/Business Intelligence) were seen as major future development. Amalgamating the results from these questions together, we believe that the knowledge areas for future development in later work packages (for example, Work Package 4) are: Data Access Technologies (Database Programmer), Database Design, Database Administration, and OLAP & Data Warehousing.

Finally, there was clear evidence of support for a vendor-independent certification scheme with 82.2% in agreement and 2.8% in disagreement.

References

ACM/IEEE. 2001. "ACM/IEEE Computing Curricula". Dec. 15 2001. http://www.computer.org/education/cc2001/ (23 July 2004).

EUCIP. 2003. *EUCIP (European Certification of Informatics Professionals) Core Syllabus*. March 2003. http://www.eucip.com/DownloadFiles/Core_Syllabus_March_2003.pdf (23 July 2004).

Appendix A - Database Technology Professional Questionnaire

Database Staffing/Database Usage

1. *What database roles does your company currently have (please tick as many as apply and indicate the approximate number currently employed in the role together with the current number of vacancies):*

	Tick	Number employed	Number of current vacancies
DBA	☐	_____	_____
Database Designer/Analyst	☐	_____	_____
Database Programmer/Developer	☐	_____	_____
Data Warehouse Manager	☐	_____	_____
Data Miner	☐	_____	_____
Others (please specify):		_____	

2. *Which DBMSs do you mainly use (please tick those that apply and indicate whether based on NT or Unix/Linux):*

	NT	Unix/Linux	Other
Oracle	☐	☐	☐
DB2	☐	☐	☐
Ingres	☐	☐	☐
Sybase	☐	☐	☐
SQL Server	☐	☐	☐
Access	☐	☐	☐
Others (please specify)	_____		

3. *Is your database application development:*

Carried out in-house ☐

Out-sourced ☐

None undertaken (mainly standard packages used) ☐

4. *Please specify which database skills/knowledge (if any) you feel are lacking in recent graduates you have hired to work on database related jobs:*

_____

_____

_____

5. *Please specify which database skills/knowledge (if any) you feel are lacking in your IT professionals working on database related jobs:*

_____

_____

_____

If you do not carry out database application development in-house, please go to Question 7.

6. Please specify which languages/tools you use for the following:

Database Backup and Recovery: _____

_____

Database Design: _____

_____

Database Development: _____

_____

Application Development: _____

_____

Future Plans

7. Do you have in production or planning any of the following developments (please tick as many as apply):

Data warehousing ☐

eCommerce ☐

Enterprise Resource Planning (ERP) ☐

and/or Customer Relationship Management (CRM) ☐

J2EE/.NET ☐

XML databases ☐

Others (please specify): _____

8. *Identify the top three (ranked from 1 to 3) skills that in your opinion are most likely to influence Database Technology (and the corresponding job market) the most in the near future:*

Object-Orientation ☐

Decision Support (OLAP, Data Warehouses) ☐

Business Intelligence (Data Mining) ☐

Web Technologies ☐

Parallel Processing ☐

Multidimensional Data Types/
Geospatial and Temporal Data ☐

Deductive Programming Logic ☐

Distribution & Heterogeneity ☐

Real-time Applications ☐

Workflow and Process Automation ☐

Intelligent Agent Technology ☐

Others (please, specify): _____

Database Certification

9. *Please specify the approximate number of staff who have the following database vendor certification:*

Oracle _____

Microsoft DBA _____

IBM _____

Others (please specify): _____

10. *Do you consider database vendor certification to be important for staff?*

Not important	Neutral	Somewhat Important	Very Important
☐	☐	☐	☐

11. Would you prefer to see a vendor-independent certification scheme for database professionals?

Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
☐	☐	☐	☐	☐

Company Information

12. Number of IT staff employed: _____   Number of current IT vacancies: _____

Total number of staff employed: _____

13. For IT companies, please identify the main 'product' your company delivers:

Database Application Development & Maintenance ☐
DBMS Vendor Certified Dealer ☐
Computer Hardware and Network Solutions ☐
Computer Software Solutions ☐
Internet Service Provider ☐
eCommerce Solutions ☐
Others (please, specify): _____

14. Business sector of your company:

Agro-Industries	Energy & Utilities	Manufacturing	Services	Construction	Public Sector	Communications
☐	☐	☐	☐	☐	☐	☐

15. Country:

Finland	Germany	Greece	Spain	UK	Other:
☐	☐	☐	☐	☐	_____

16. Company name (optional): **_____**

Please tick if your office is a branch of a larger organization: ☐

Thank-you for your participation in this survey.